AD_____

Award Number: DAMD17-03-1-0697

TITLE: Computerized Identification of Normal Mammograms

PRINCIPAL INVESTIGATOR: Robert Nishikawa, Ph.D.

CONTRACTING ORGANIZATION: University of Chicago
Chicago, IL 60637

REPORT DATE: October 2005

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-10-2005	Final	30 Sep 2003 – 29 Sep 2005

4. TITLE AND SUBTITLE
Computerized Identification of Normal Mammograms

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-03-1-0697

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Robert Nishikawa, Ph.D.

E-mail: yue@bme.ri.ccf.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Chicago
Chicago, IL 60637

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
Original contains color plates: All DTIC reproductions will be in black and white.

14. ABSTRACT
The purpose of this concept-award project is to develop an automated method to identify normal mammograms, that is those without breast disease. This is a new paradigm in computer-aided diagnosis (CAD), since all other CAD schemes identify breast cancer. We are relying on the natural pattern of glandular tissue in the normal breast, which radiates out from the nipple. Breast cancer disturbs this pattern. We have developed a database of approximately 20,000 regions of interest (ROIs) of normal breast tissue and 100 regions containing a portion of a breast cancer. Each region was automatically extracted from a mammogram that was reduced in size. These ROIs were used to train an artificial neural network called a self-organizing map (SOM) to learn the mammographic pattern of normal breast tissue. SOMs are self-learning classifiers that categorize input data into a number of distinct classes. The SOM was able to train to classify normal and abnormal ROIs. However, to date our method correctly identified only 10% of the normal cases as normal, and 92% abnormal cases as not normal. Our target was to achieve between 25% and 50% recognition of normal mammograms. As a result of this study, several new ideas were generated that could improve the performance of the technique.

15. SUBJECT TERMS
mammography, detection, computer-aided detection, normal

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	UU	16	

Table of Contents

4. INTRODUCTION

Computer-aided detection (CADe) systems have sensitivities at least equal to radiologists, 80-90% depending on the system, but the false detection rate is more than a magnitude higher than that of radiologists (on average the computer has 2 false detections per case, whereas a radiologists will have a one false positive every 10 cases). Because of the high false detection rate, a radiologist must review virtually every mammogram. Instead of locating abnormalities in mammograms, as is done with all current CADe systems, we propose to develop a method for determining normal mammograms. Initially, our approach would allow the radiologist to read only those cases that are judged to be not normal, reducing the number of cases reviewed potentially to 90% or better, allowing for more time to read cases that are more likely to contain a malignancy. Ultimately, if our approach is effective and optimized, it could be used as a front-end (triage system) to conventional CAD schemes that could be optimized to run on the "not normal" cases. Furthermore, we believe that the ultimate performance of CAD systems will not improve to the level of a radiologist using the current paradigm. A normal breast has a pattern of structures radiating out from the nipple. A cancer can disrupt this pattern. Our approach is to use this radiating pattern as a basis for recognizing normal mammograms. We will process the image to highlight the radiating pattern. Then by taking small regions of interest (ROIs), we will train a classifier to recognize normal ROIs. The classifier used in this study is a specialized artificial neural network called a self-organizing map (SOM) {1}.

5. BODY

5.1. Tasks

Task 1. Process image to highlight ductal system:

a. Assemble 2,000 consecutive digitized normal screening exams and 100 cancer exams (cc views only) from an existing database of 25,000 consecutive screening mammograms.

b. Create 3 datasets: (i) development set (500 normals); (ii) training set (1000 normals and 75 cancers); and (iii) testing set (500 normals and 25 cancers).

c. Reduce image size by a factor of 10, testing different methods such as mean, maximum, median, and rank order.

d. Implement two processing techniques, morphological operators and a linear detection algorithm developed by Zwiggelaar *et al*. (using the development dataset)

Task 2. Train support vector machine to recognize normal mammograms:

a. Train support vector machine (using the training dataset)

b. Measure the performance of the technique (using the testing dataset)

5.1.a Assemble databases

In a previous 5-year project, we digitized over 20,000 consecutive screen-film

mammograms to 10 bits and 100-micron pixel size {2}. From this dataset, we have assembled 54 cancer-free consecutive cases and 12 cancer cases, collecting only the cranio-caudal (cc) views. The abnormal cases contain a mass that was biopsied and found to be malignant. Approximately half of the abnormal cases contained a cancer that was initially missed clinically, but could be seen retrospectively once the cancer was discovered a year or more later. The normal cases were obtained by reading all the radiology reports for that patient. In a separate process, these reports had all patient identifiers removed and all reports from a single patient were placed in a single file and identified by the study number that was generated previously to allow the radiology report to be associated with the image. The study number is not traceable to any patient identifier. The mammograms are devoid of patient identifiers. To be considered normal, the case must have had at least a two-year period in which the mammograms were considered normal. Further, we selected from these cases, cases that were free of any type of lesion, including obvious benign findings such as lymph nodes and calcified vessels. This subset was used in the development data set.

In the development phase, the goal was to understand how to pre-process the image and to understand how the SOM works. To do this, we needed only a small database with very few cancer cases, in part because a large number of regions-of-interest (ROIs) can be selected from each image. We had planned to enlarge the database, but we had difficulty in using the SOM and therefore we devoted all our effort to implementing the SOM. As a result, we did not finish collecting all the cases. We will enlarge the database in the near future to train and test our method more thoroughly.

5.1.b. Preprocess the mammograms

The 54 normal cases and the 12 abnormal cases were preprocessed to produce ROIs that either contain a portion of a cancer or are cancer free. This was done in four steps.

Step 1. The breast border was determined using software previously developed in our laboratory {3}.

Step 2. Wavelet decomposition was applied to the image using a bi-orthogonal spline mother wavelet implemented in MATLAB. All mother wavelets available in MATLAB were tested, but the bi-orthogonal spline gave the best visual result. This mother wavelet was used by Strickland in his study of detecting mammographic calcifications using wavelets {4}. We constructed the magnitude image from the horizontal and vertical components of the wavelet transform using level 3 (see Fig. 1). We originally had planned to implement a morphological operator and a linear detection algorithm developed by Zwiggelaar *et al.*{5} We spent several weeks implementing the morphological operator method but it did not produce satisfactory results. To save time, we implemented the wavelet filtering method in MATLAB.

Step 3. Based on the estimated breast border the largest rectangle that fit in the breast boundary was extracted from the wavelet image. From this rectangle, overlapping candidate ROIs that were 128x128 pixels in size were extracted. Each candidate ROI was shifted by 64 pixels from the previous candidate ROI. For each candidate ROI, a histogram of its pixel values was calculated. An upper and lower bound threshold was used to filter out “partial ROIs” (i.e., those that include non-breast tissue). Partial ROIs had either a substantial number of pixels that were white (e.g., if a metallic marker was

present) or black (e.g., if the estimated breast border included some non-breast area). From the 108 normal mammograms (two views from each case) there were a total of 20,679 ROIs selected, or approximately 200 per image. From the 12 abnormal cases, 102 ROIs were selected and each ROI contained a portion of the breast cancer that presented as a mass.

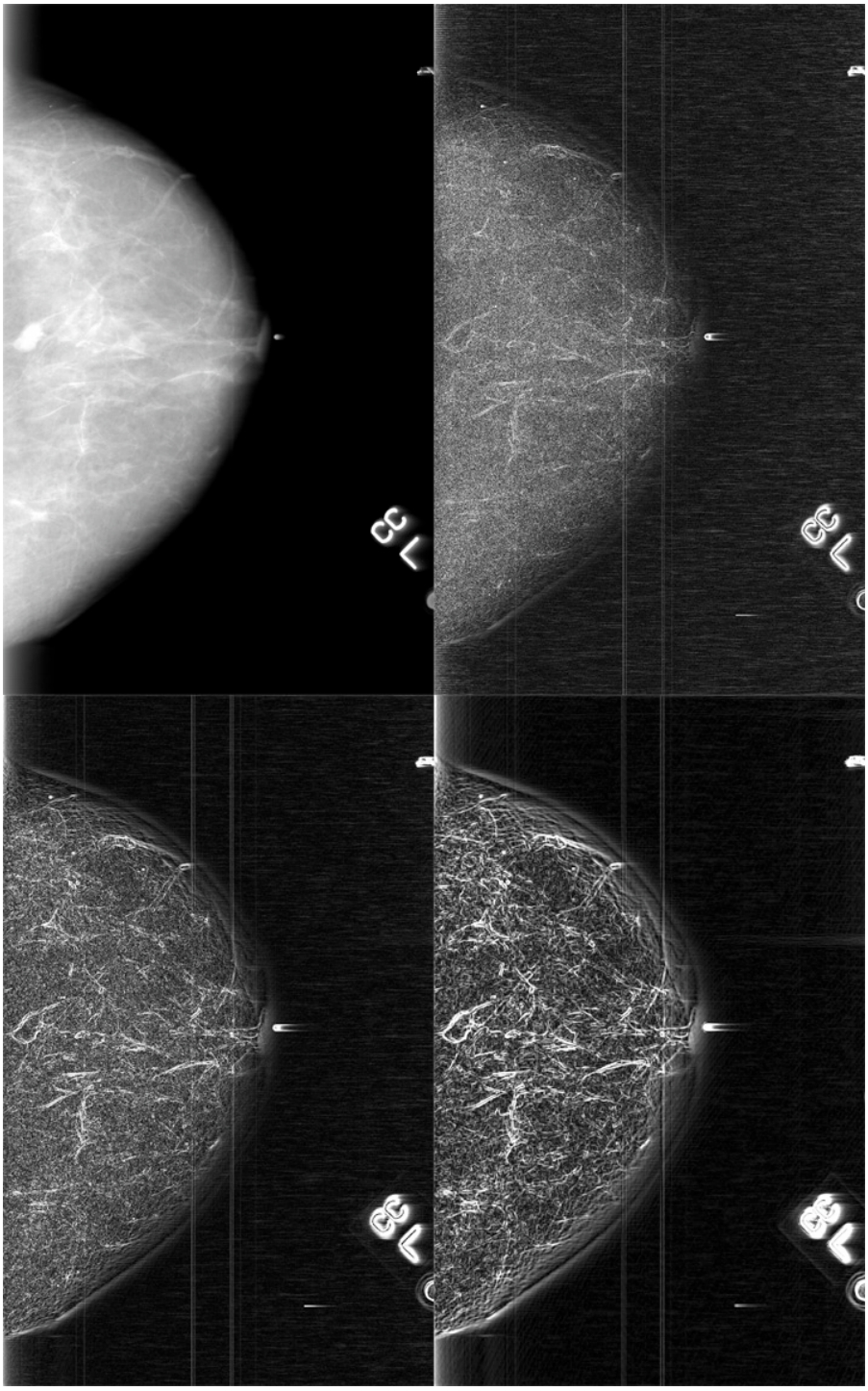

Figure 1. Illustration of the wavelet preprocessing. The original image is shown in the upper left. The other three images are the magnitude image of the wavelet transform for

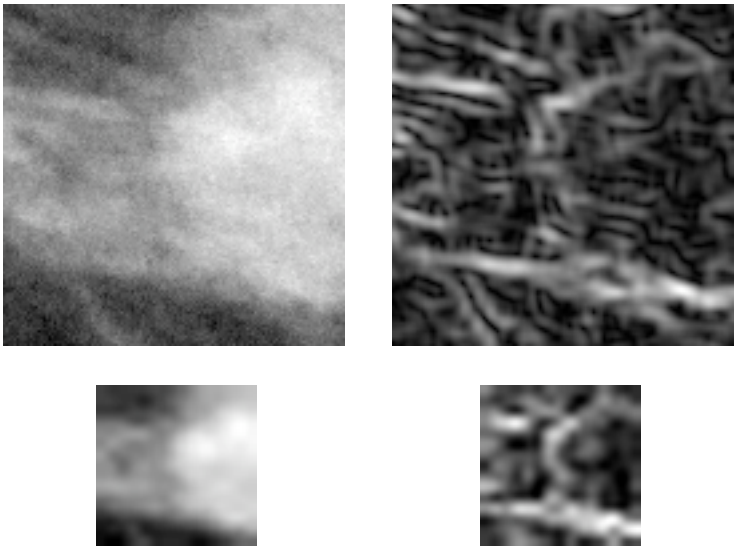

Figure 2. An illustration of the down sampling of the regions-of-interest (ROI). The top row show a 128x128 ROI extracted from the original image (left) and the wavelet processed image (right). The bottom row shows the two images after 8x8 pixel averaging. These two ROIs have been enlarged by a factor of 4.

level 1 (upper right), level 2 (lower left) and level 3 (lower right). We used level 3 in this study.

Step 4. Each ROI was then reduced in size by averaging 8x8 pixels together. This produced a 15x15 pixel ROI. (One row and one column were lost in MATLAB average subsample algorithm for some unknown reason.) This produced good results visually (see Fig. 2), so no other down sampling methods were tried. We plan in the future to use a median down sampling method to see if we get improved results.

5.2. Train classifier to recognize normal mammograms

In our original statement of work, we proposed using a support vector machine (SVM) as our classifier {6}. We have however, decided to us a self-organizing map (SOM) for the following reasons {1}.

1. There are many different appearances of breast lesions (e.g., calcifications, circumscribed masses, spiculated masses, etc.). There are even more different appearances of normal breast tissues, since the appearance of normal breast tissues depend upon breast thickness, breast density, amount of breast compression, the parenchymal (Wolfe) pattern, position in the breast, etc. Given the wide variety of both normal and abnormal patterns, it would take a very sophisticated (or complex) classifier to classify all possible normal and abnormal breast patterns into two classes. SVMs are designed to produce two classes, while SOMs are designed to handle multiple classes.

2. SOM is an unsupervised classifier and SVM is a supervised classifier. The important

difference is that for supervised classifiers, one needs to know the classes in the problem. Even if one decided to use multiple classes with a SVM, the classes must be defined *a priori*. However, we do not know *a priori* all the possible different classes. We believe that an unsupervised classifier is ideally suited to this problem, as it will determine the number of classes present in the data.

3. An SVM relies on data that are on the "border" between the two classes. Since most normal patterns are very different from abnormal patterns, any training example that is obviously normal will not be "useful" for training. In this problem, most of the normal training examples will not be useful. An SOM relies on all training samples.

An SOM is useful for reducing multi-dimensional data – 225 (15x15) dimensions in our study – to a two-dimensional surface. An SOM consists of a 2-D array of nodes. Each node represents a category based on a 225-element vector – each element corresponds to pixel value in a pixel in the ROI. This vector is the weights of the SOM. When trained, the SOM adjusts the vector at each node to best match the training data. The first training ROI is compared to each vector at all the nodes. The node that has a vector most similar to the ROI is selected and its vector and those in a neighborhood surrounding the select node are adjust to be more similar to the input ROI. This is repeated for each ROI in the training set, after which one training epoch has been completed. Multiple iterations or epochs are used to train the SOM. After training, given an input ROI, the SOM will output which node or category that ROI belongs, so the output of the SOM is a number between 1 and the number of nodes.

The implementation of the SOM we used was the SOM Toolbox 2.0 software library running in MATLAB™ (version 6.1.0 and 7.0 by Mathworks, Inc.). We originally used the nnet toolbox from MATLAB™ but we were unable to get satisfactory results with this implementation of the SOM. Further, the SOM Toolbox ran approximately 10 times faster than nnet.

5.2.a. Train Classifier

Since we did not have experience using SOMs, we first did some preliminary studies to test the reliability of the SOM for our problem. After considerable effort, described in part in our last annual report, we were able to train the SOM to get reproducible results. This ultimately required that we use a different implementation of the SOM software (see previous paragraph). However, we needed to try many different approaches, including using unprocessed ROIs, instead of wavelet-processed images. Thus for the remainder of the project, we used unprocessed ROIs. We plan to implement the SOM using processed ROIs in the future.

The SOM was able to self-determine the size of the network and used a 12x52 set of neurons or nodes in a hexagonal grid pattern. Each node corresponds to one class of the ROIs so that 624 different patterns can be classified. The size was calculated based on the number of data samples (n), where the total number neurons (m) was defined by a heuristic formula of $m = 5\sqrt{n}$. The ratio of the side lengths was based on the ratio between the two biggest eigenvalues of the covariance matrix of the given data, and the actual side lengths were then set so that their product was as close to the desired neuron number (m) as possible.

For training we used 15,719 normal ROIs and 49 abnormal ROIs. To get better balance between the number of normal and abnormal ROIs, each abnormal ROI was replicated ten times

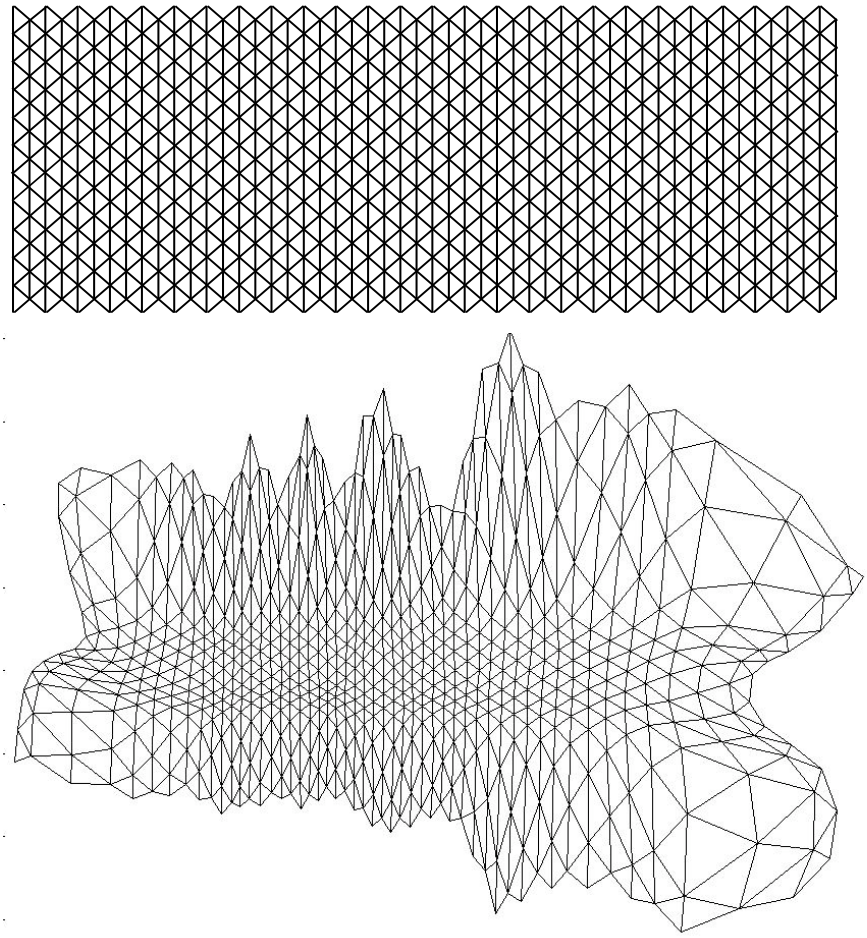

Figure 3. The topology of the untrained SOM is shown at the top. At the bottom is the trained SOM. The intersection of lines indicates the position of the nodes, which are initially in a 12x52 hexagonal pattern.

in the training set. In our original approach we were only going to train the SOM using normal ROIs. Our hope was that abnormal ROIs would be classified to nodes that were sparsely populated with normal ROIs. This was true for most abnormal ROIs, but there were some that were classified to well-populated nodes. As a result, this approach did not work. We are currently using both normal and abnormal ROIs to train the SOM. In this way, abnormal ROIs can be placed in nodes that reflect the appearance of abnormal ROIs.

The trained SOM map is shown in Fig. 3. The initial state of the SOM is a regular grid of equally spaced nodes. The trained SOM shows a different topology. Nodes that are close to each other indicate well-populated nodes (center of the SOM map), while sparsely populated nodes are spaced further apart from other nodes (e.g., the right side of the SOM map). Each node is associated with a set of weights that are a 15x15 array (corresponding to the size of the input ROIs). The weights essentially form a template of different patterns in the input ROIs. Therefore, one can associate with each node, the pattern of the ROI that is classified to that node. Figure 4 shows the weights for each node in a rectangular grid corresponding to the hexagonal grid of the SOM. (Conceptually, a rectangular grid can be transformed into a hexagonal grid by shifting ever other row by half a space.) From Fig. 4 one can see a general pattern. From left to

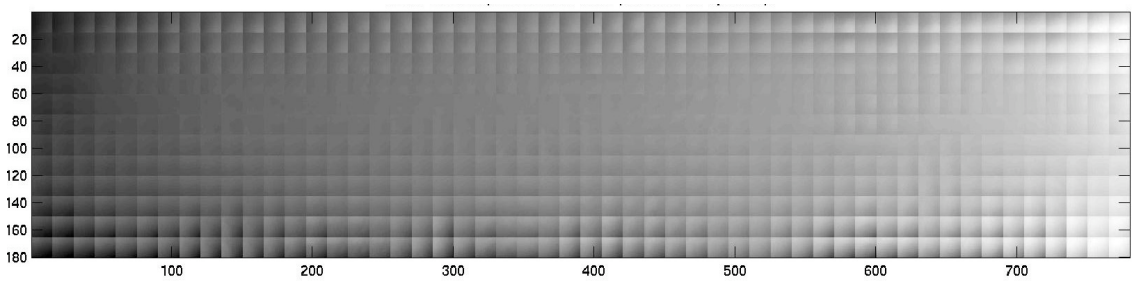

Figure 4. A "map" of the trained nodes of the SOM. There are 12x52 squares each corresponding to the weights of one node. Each node has 15x15 weights. As the SOM is trained, different patterns that appear within different ROI are classified to different nodes. Like patterns are classified to the same node. Similar patterns are classified to neighboring nodes. The weights are essentially used as a template and compared to the pattern present in an ROI. ROIs are classified to the node whose weight template most closely matches the pattern within the ROI. As the SOM trains the weights are adjusted to best match the training ROIs.

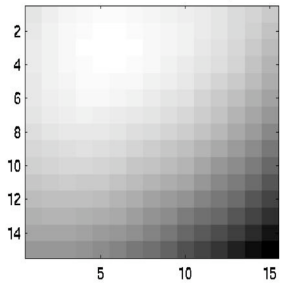

Figure 5. An enlargement of node centered at 560,175 (i.e., bottom row approximately midway between 500 and 600 in Figure 4). The grey-scale corresponds to the value of one of the weights of the trained SOM. Each node consists of 15x15 weights.

right, the patterns go from dark to light. Further, by examining Fig 3, abnormal ROIs are sparsely populating the right side of the SOM corresponding to having a bright ROI, which is expected, since cancers appear bright in an image. An enlargement of one of the nodes in Fig. 4 is shown in Figure 5.

5.2.b. Measure the performance of the technique

The output of the SOM for a given ROI is a number that is the node to which the ROI was classified. Figure 6 shows to which nodes ROIs from a normal mammogram are classified. Figure 7 shows the same for all the ROIs from actually abnormal mammograms and just the abnormal ROIs from the actually abnormal mammograms.

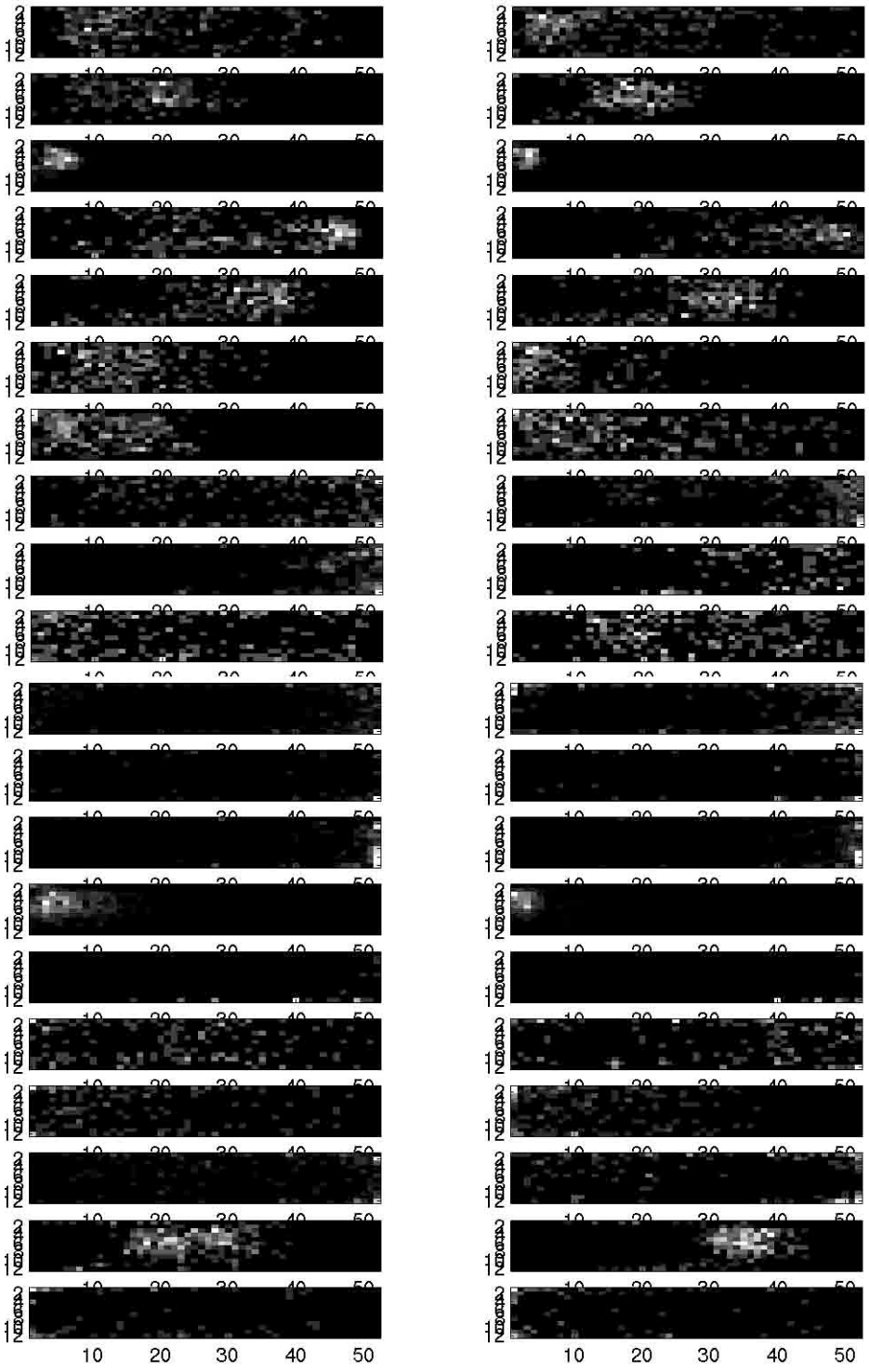

Figure 6. SOM output maps for normal cases. The left column is for the left breast mammogram and the right column is for the right mammogram. Each row is for a different case.

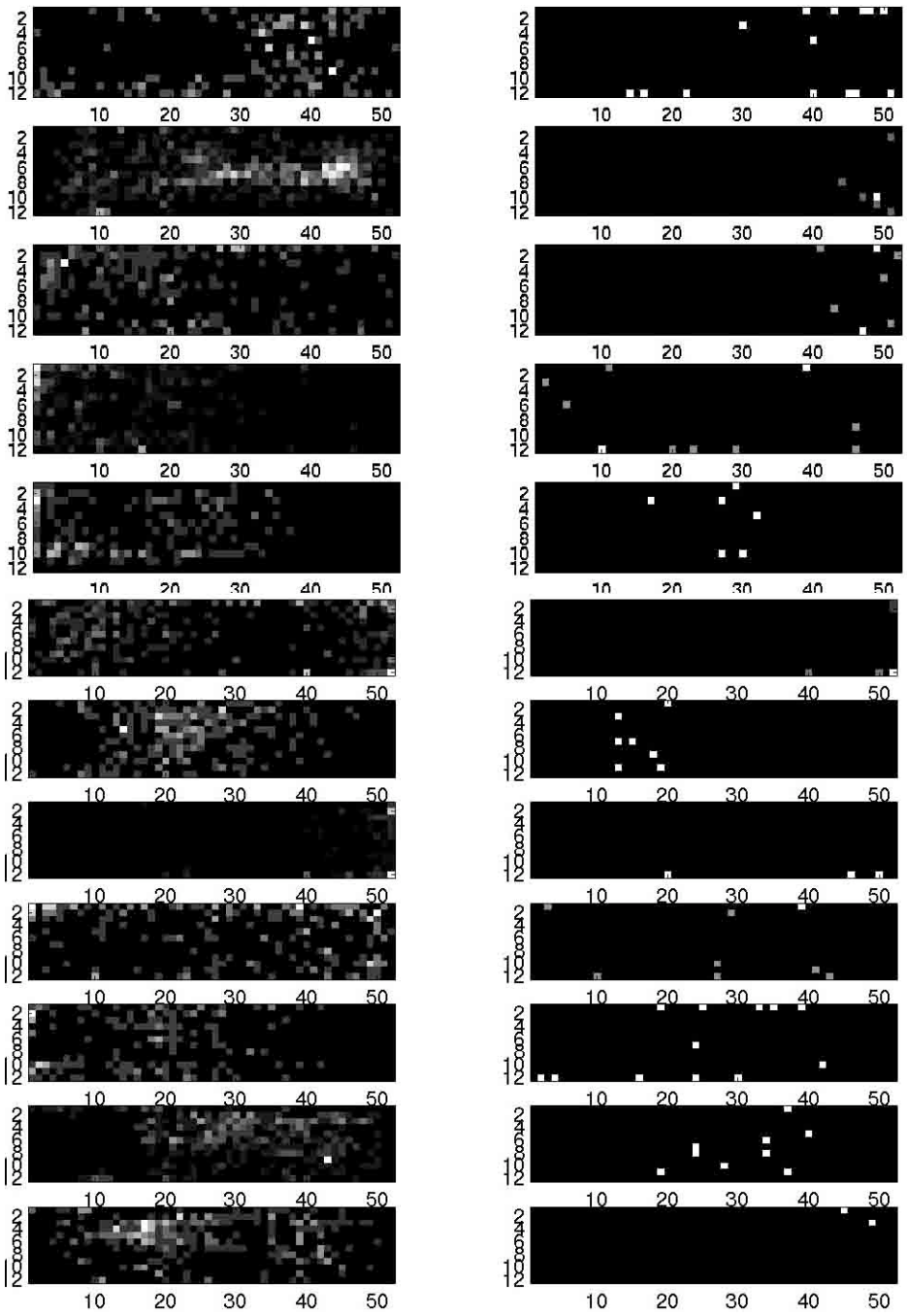

Figure 7. SOM output map for abnormal images. The left column is for all ROIs in an image that contains a cancer. The right column is for only ROIs that contain a piece of the cancer. Each row is for a different case.

The goal of the project is to classify cases as either normal or not normal (as opposed to normal and abnormal). When implemented clinically, the SOM would triage screening mammograms to definitely normal and can't rule out the possibility that the woman may have cancer. The radiologists would only read the latter group.

There are many different ways to analyze the output of the SOM to determine whether an image is normal or not normal. We chose the following method. For each case, each mammogram is divided into multiple ROIs, as described in Section 5.1.b. These ROIs are analyzed by the SOM and a "map" (i.e. a two-dimensional histogram, where the number in each bin of the histogram is encoded as a gray level in the map) of distribution of the node to which the ROIs were classified (see Figs. 6 and 7). We also determine the map for all abnormal ROIs (those which contain a piece of a cancer) from a given abnormal image (see Fig. 7). The maps for abnormal ROIs differ from some of the maps from normal cases (Fig. 6), with the abnormal ROI maps tending to be predominantly in the right part of the map and some normal cases being in the left side of the map.

By setting a threshold on the x-axis of the map, we can classify cases as normal or not normal. For example, the lowest maximum node value along the x-axis for abnormal ROIs is 20. That is, for the dataset we used, we can correctly classify all actually abnormal cases as not normal if the case has at least one ROI above 20. Conversely, cases that do not have any ROIs above 20 are classified as a normal case. Table 1 gives the performance for the classifier based on this approach and Fig. 8 gives the graphical representation in terms of percentages. For not including any actually abnormal cases as normal 4% of the actual normal cases can be identified as

Table 1. The performance of our technique for identifying normal mammograms. The threshold value is applied to the x-axis of the SOM maps (see Figs. 6 and 7). If there is a node that is populated and it is less than or equal to the threshold value then the image is considered not normal.

Threshold Value	Number of Actual Abnormal Cases Called Abnormal	Number of Actual Normal Cases Called Normal
1	12	0
19	12	2
31	11	6
39	10	10
41	9	10
42	8	10
45	7	12
48	6	18
49	5	19
50	4	25
51	2	28

normal. However, radiologists do not have 100% sensitivity. For misclassifying one actual abnormal case (sensitivity= 92%), 11% of the normal cases can be classified as normal and for 2 misclassified abnormal cases (sensitivity= 83%), 19% of the normal cases can be correctly classified.

We note that this result is optimistically biased because all cases were used in the evaluation and a fraction of them were used in training the SOM. We also note that the performance of the SOM is probably suboptimal because of the small number of abnormal cases (n=5) used in training. In the future, when we enlarge our database, we will create independent datasets for training and testing.

Based on discussions with radiologists in our department, our minimum goal is to identify 25% of normal cases at 95% sensitivity. Although we did not reach that goal, we believe that with further research that goal is obtainable. Future approaches for improving our method is described in Section 5.3.

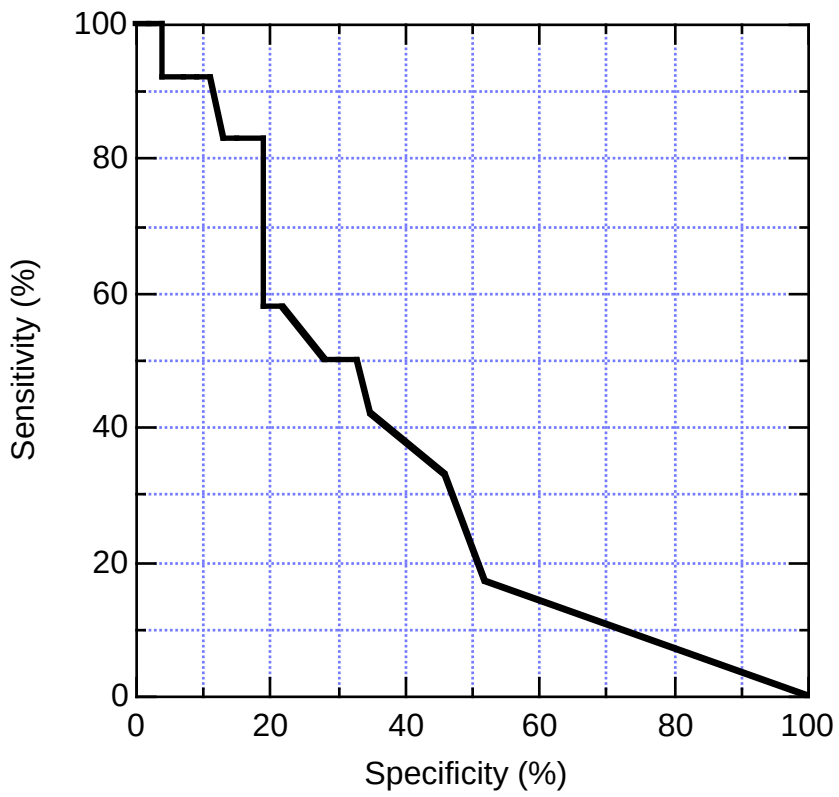

Figure 8. A graph of sensitivity versus specificity for our technique for identifying normal mammograms. The specificity is the percentage of actual normal cases call normal.

5.3 Recommendations in relation to the Statement of Work

We implemented two changes to our original statement of work. First, we preprocessed the images using a wavelet filter instead of two methods proposed: a morphological operator and a linear detection algorithm developed by Zwiggelaar *et al*. This was done because we could not get the former method to work effectively and the wavelet method was faster to implement. Second, we used a self-organizing map (SOM) classifier instead of a support vector machine (SVM). The reason for this change is given in Section 5.2.

Unfortunately, we did not complete all the tasks proposed sufficiently to make any definitive conclusions. We believe that we have a good framework upon which to continue developing our technique. As a result of the completed research, we will do the following studies:

1. Preprocess the ROIs using wavelets, as described in Section 5.1.b, use the preprocessed ROIs as input to the SOM.

2. Normalize the mean pixel value in each ROI, instead of using wavelets. Currently, we believe that the SOM is putting to much emphasis on the mean pixel value and less emphasis on the structure within the ROI. By ensuring that each ROI has the same mean pixel value, we believe that the SOM can more accurately classify normal and abnormal ROIs.

3. Use a larger dataset for training and testing of the technique.

4. Examine different methods for reducing the image size (e.g., selecting median value instead of the mean).

6. KEY RESEARCH ACCOMPLISHMENTS

- Developed technique for identifying normal mammograms
- Developed a method for reducing image size and preprocessing the images
- Trained SOM to classify normal and abnormal ROIs

7. REPORTABLE OUTCOMES

The initial technique can identify 11% of normal cases at sensitivity of 92%. We expect that with more cases to train and test the technique better results can be obtained.

8. CONCLUSIONS

We have developed a method for the automated identification of normal mammograms. The technique uses a self-organizing map (SOM), which is a special unsupervised artificial neural network that is capable of learning normal and abnormal mammographic patterns.

Preliminary results indicate the potential of the technique – 11% of the normals correctly identified at a sensitivity of 92%. We believe that with further research, as outlined in this report, we will be able to develop a technique that can correctly identify at least 25% of normal cases at sensitivity of 95%.

9. REFERENCES

1. Kohonen T: *Self-Organizing Maps* (Springer-Verlag, New York, 1997).

2. Nishikawa RM, Giger ML, Schmidt RA, Wolverton DE, and Doi K: Prospective testing of a clinical CAD workstation for the detection of breast lesions on mammograms. In: *Computer Aided Diagnosis in Medical Imaging*. Doi K, MacMahon H, Giger ML, and Hoffmann KR, Eds. (Elsevier, Amsterdam, 1999), pp. 209-214.

3. Bick U, Giger ML, Schmidt RA, Nishikawa RM, Wolverton DE, Lu P, Vyborny CJ, and Doi K: Automated segmentation of digitized mammograms. *Academic Radiology* 2: 1-9, 1995.

4. Strickland RN, and Hahn H: Wavelet transforms for detecting microcalcifications in mammograms. *IEEE Transactions on Medical Imaging* 15: 218-229, 1996.

5. Zwiggelaar R, Parr TC, Schumm JE, Hutt IW, Taylor CJ, Astley SM, and Boggis CR: Model-based detection of spiculated lesions in mammograms. *Med Image Anal* 3: 39-62, 1999.

6. El-Naqa I, Yang Y, Wernick MN, Galatsanos NP, and Nishikawa RM: A support vector machine approach for detection of microcalcifications. *IEEE Transactions on Medical Imaging* 21: 1552-1563, 2002.

10. APPENDICES

None